



COVID-19: data-driven dynamics, statistical and distributed delay models, and observations

Xianbo Liu · Xie Zheng · Balakumar Balachandran 

Received: 25 May 2020 / Accepted: 29 July 2020 / Published online: 6 August 2020
© Springer Nature B.V. 2020

Abstract COVID-19 was declared as a pandemic by the World Health Organization on March 11, 2020. Here, the dynamics of this epidemic is studied by using a generalized logistic function model and extended compartmental models with and without delays. For a chosen population, it is shown as to how forecasting may be done on the spreading of the infection by using a generalized logistic function model, which can be interpreted as a basic compartmental model. In an extended compartmental model, which is a modified form of the SEIQR model, the population is divided into susceptible, exposed, infectious, quarantined, and removed (recovered or dead) compartments, and a set of delay integral equations is used to describe the system dynamics. Time-varying infection rates are allowed in the model to capture the responses to control measures taken, and distributed delay distributions are used to capture variability in individual responses to an infection. The constructed extended compartmental model is a nonlinear dynamical system with distributed

delays and time-varying parameters. The critical role of data is elucidated, and it is discussed as to how the compartmental model can be used to capture responses to various measures including quarantining. Data for different parts of the world are considered, and comparisons are also made in terms of the reproductive number. The obtained results can be useful for furthering the understanding of disease dynamics as well as for planning purposes.

Keywords Dynamics and control of epidemics · Generalized logistic function · System identification · SEIQR model · Delay integral equations

1 Introduction

Since the first reported case of novel coronavirus (SARS-CoV-2) in Wuhan, China, toward the end of 2019, this highly infectious disease first spread rapidly within China and to its neighboring countries [1, 2]. After China, the next confirmed cases occurred in Japan, South Korea [3], and Thailand in late January, and soon thereafter in 2020, the USA reported its first case in Washington State. In February, Europe faced its first outbreak in Italy [4], with churches and schools closing immediately thereafter and towns being locked down [5]. By early March, following the World Health Organization's declaration of the coronavirus as a pandemic, the USA imposed a ban on travelers from Europe and declared a national emergency on

X. Liu
State Key Laboratory of Mechanical System and Vibration,
School of Mechanical Engineering, Shanghai Jiao Tong
University, Shanghai 200240, China
e-mail: liuxianbo@sjtu.edu.cn

X. Liu · X. Zheng · B. Balachandran (✉)
Department of Mechanical Engineering, University of
Maryland, College Park, MD 20742, USA
e-mail: balab@umd.edu

X. Zheng
e-mail: xiezheng@umd.edu

March 13. As of May 4, more than 253,000 people had died from COVID-19, with 3.6 million infections confirmed in more than 180 countries and territories globally [6]. The USA has taken the hardest hit from the pandemic, with 1.22 million confirmed cases and more than 70,000 deaths, at the time of writing of this paper. The goal of this work has been to look at the data available on the number of infections and study the evolution of the infection dynamics.

In order to facilitate the authors' quest to pursue data-driven dynamics based on infection data, a statistical approach based on the generalized logistic function has been taken along with studies of a delay differential system. The logistic function, which was introduced by Pierre Verhulst for population growth modeling [7], is now widely used in various areas of science and engineering. Applications include friction modeling in mechanical systems [8], activation function in neural networks [9], and infectious disease spreading in biological systems [10]. The generalized logistic function, which has an asymmetric form in between the lower and upper horizontal asymptotes, was used by Richards [11] for modeling plant growth. This function has been recently used for studying disease dynamics [12]. As for the studies in spreading dynamics, logistic functions and generalized logistic functions are essentially compartmental models [13] with a susceptible state and an infected state (SI model). In reality, there exists a latent period, when people are infected but not yet infectious. Consideration of this aspect leads to the SEIR model [14], in which one has susceptible, exposed, infectious, and removed (recovered or dead) states. Moreover, if mitigation measures are applied, a new state of quarantine needs to be considered, which results in the SEIQR model. However, in all of the aforementioned models, the derivatives of the different states are only dependent on their current values and assumed to be uniformly distributed. Here, these models are extended through introduction of distributed time delays.

Delay differential equations (DDEs) arise often in various science and engineering applications, for instance, in mechanical engineering [15–17]. Different from ordinary differential equations, to solve DDEs, one needs information on current states and past states over time intervals in the past [18]. DDEs are critical for modeling the spreading of COVID-19, since time delays can be used to capture the durations of the latent, quarantine, and recovery periods. A SEIR model with two constant delays can be found in the work of [19],

a study of the global behavior of SIER with delay can be found in the studies [20, 21], and an SIR model with delays has been studied in [22]. Furthermore, stability studies on the solutions of an SIR model can be found in [23, 24], and solutions of an SEIR model can be found in the work of [25].

A primary objective of this paper is to use data-driven dynamics to further the current understanding of key aspects and features associated with the COVID-19 outbreak, as well as to help assess the viability of control strategies that are applicable for mitigation, for example “flattening the curve.” The remainder of the paper is organized as follows. In Sect. 2, based on the susceptible-infected (SI) populations model, the S-shaped logistic and generalized logistic functions are introduced to describe the outbreak of the COVID-19 among various countries—China, Italy, Germany, and the USA. With the goal of finding the inflection points for these countries, parameters of the S-curve functions for different countries are identified using an optimization method. In Sect. 3, by considering the incubation period and the quarantine state, an improved SEIQR model with distributed time delays is developed. The time delay is employed here to capture the time gap between different (compartmental) states in the new model. With data of confirmed cases, parameters are identified and simulations together with prediction of different geographic regions are presented. With the developed models, potential mitigation strategies for different scenarios are proposed. Finally, concluding remarks are provided.

2 Generalized logistic function based data-driven dynamics of COVID-19

In this section, the infection data of COVID-19 for the US, as well as different countries from all over the world, are gathered by using a web-read function of MATLAB from the website of The COVID Tracking Project [26] and Worldometers [27], respectively. The determination of unknown parameters in the generalized logistic function is driven by the data of the total confirmed number of positive cases of the COVID-19, and a nonlinear regression algorithm is used. Then, the inflection (peak) point of COVID-19 is predicted and analyzed for different countries and regions. Followed by that, considering each of the regions as an element of the overall global system, a composite global model

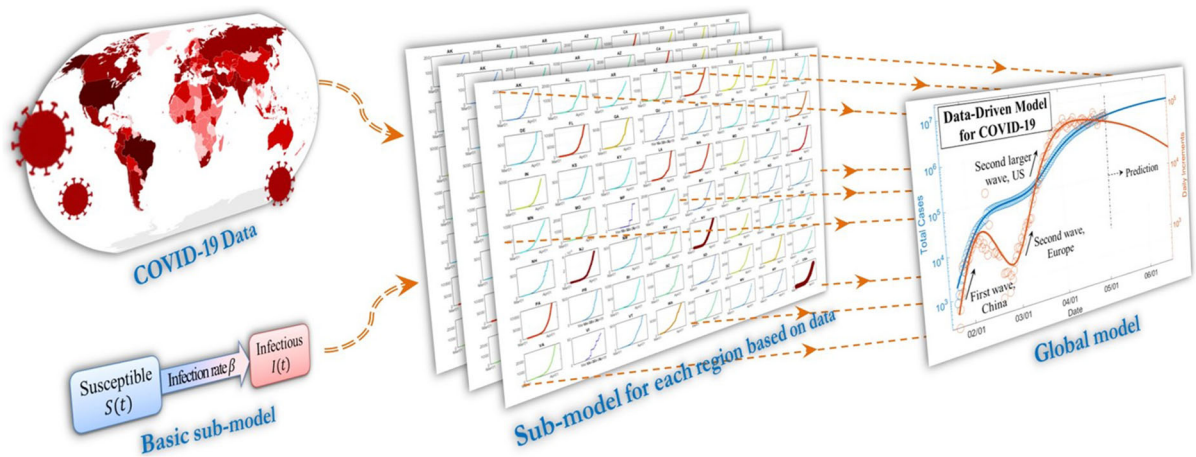


Fig. 1 Development of the composite global model and data-driven dynamics

which is comprised of 148 sub-models is developed, for capturing and tracking the global COVID-19 dynamics, as shown in Fig. 1.

2.1 The generalized logistic function

The susceptible-infectious (SI) model, which is also called the simple epidemic model (SEM), is the simplest form of all epidemic models, as shown in Fig. 2a. The evolution of the infection number in an SI model is also known as logistic growth and results in a *logistic function* (sigmoid function). The logistic function is centrosymmetric about the inflection point, which means the decreasing (saturation) period mirrors the growing period. However, for the COVID-19, from the current data, it can be gleaned that the infection growth occurs as a short outburst, followed by a long saturation period. In this scenario, the logistic function is no longer an appropriate descriptive function, since the saturation period needs to be slowed down. The *generalized logistic function*, also known as Richards' curve, originally developed for plant growth modeling, is an extension of the logistic or sigmoid functions, allowing for more flexible S-shaped curves. By varying only one parameter in the generalized logistic function [11], the saturation period of the pandemic can be controlled, which makes the model and prediction easier and more realistic. Hence, based on current data, the authors choose the generalized logistic function here to capture the infection growth dynamics of COVID-19. The generalized logistic function is the solution of

the Richards' differential equation (RDE), which is a nonlinear dynamical system, can be written as follows

$$\frac{dI}{dt} = \beta I \frac{S}{N} = \beta I \left(1 - \left(\frac{I}{N} \right)^\nu \right) \tag{1}$$

Then, the generalized logistic function is obtained via direct integral as

$$I(t) = \frac{N}{\left(1 + e^{\beta \nu (t-t_0)} \right)^{\frac{1}{\nu}}} \tag{2}$$

In Eqs. (1), (2), $S = S(t)$ is the number of susceptible individuals in a population while $I = I(t)$ is the number of infectious individuals. $N = S(t) + I(t)$ is a constant, which typically represents the total population. The infection rate $\beta > 0$ represents the rate of spread of a pandemic and is the probability of transmitting disease from an infected individual to a susceptible individual. $\nu > 0$ is the parameter affecting the position of inflection point as well as the saturation period. For $\nu \equiv 1$, one has a symmetrical S-curve about the inflection point, that is, the logistic function. For $\nu > 1$, the saturation period is shorter and the result is a faster saturation compared to that noted with the logistic function. For $0 < \nu < 1$, the saturation period is longer than that obtained with the logistic function. Therefore, for the COVID-19 outbreak, it is expected that with the generalized logistic function and an optimized $0 < \nu < 1$, one will have a better match with the data. As shown in Fig. 2b, the circles represent the number of confirmed positive cases in the USA from March 1, 2020, to May

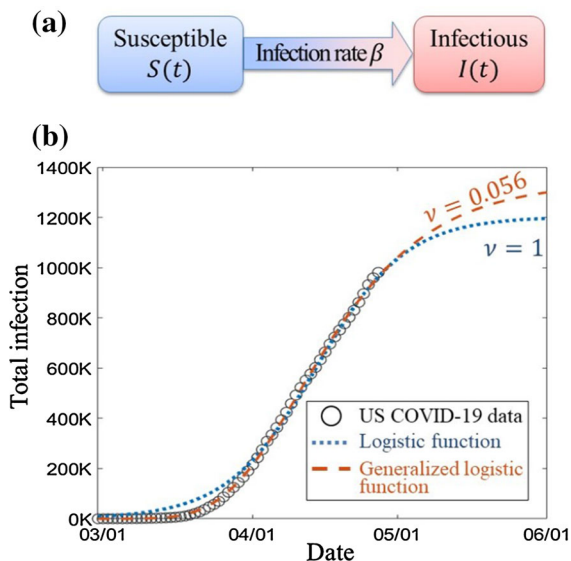


Fig. 2 a SI model and b comparison between statistical fits with the logistic function and the generalized logistic function for the US COVID-19 data

4, 2020. The dotted line in blue is the least-squares fitting done by using the logistic function ($\nu = 1$), while the dashed line in orange represents the least-squares fitting carried out by using the generalized logistic function with an optimized $\nu = 0.011$. The goodness of fit of the models for the COVID-19 data shows that the root mean squared error (standard error) of the generalized logistic function is 1.15×10^4 , which is much smaller than that for the logistic function (3.7×10^4). The generalized logistic function provides a better fit because it is both unbiased at the growth-period and saturation-period and produces smaller residuals. Four parameters $L, k, t,$ and ν in the generalized logistic function need to be identified from the COVID-19 data, by using nonlinear regression algorithms.

2.2 Infection evolution: identified dynamics and inflection point of COVID-19

The generalized logistic function Eq. (1) has four unknown parameters $L, k, t,$ and ν , which should be identified and optimized based on the COVID-19 data. The nonlinear regression algorithm adopted here is a form of regression analysis by which the data can be modeled as any nonlinear function, such as the generalized logistic function. The objective of the nonlinear regression algorithm is to find an optimized point in

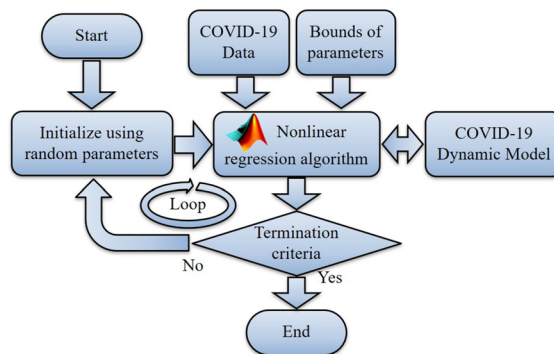


Fig. 3 Diagram for parameter identification, based on the nonlinear regression algorithm in MATLAB

the four-dimensional parameter $L - k - t - \nu$ space, through minimization of the root mean squared error (standard error) between the COVID-19 data and the model predictions.

Although both the linear and nonlinear regression algorithms are integrated into the Curve Fitting Toolbox of MATLAB R2018a, finding the global optimal point in the four-dimensional parameter space is still not an easy problem due to the nonlinearity of the generalized logistic function. Therefore, to avoid a local optimum, the parameter identification processes with Eq. (1) are carried out as shown in Fig. 3. For each loop as illustrated in the figure, the nonlinear regression algorithm is used to find a local minimum starting from widely varying initial values (random) of the parameters. And finally, the most extreme one is chosen from these local minima as the global minimum when the termination criteria are satisfied. Here, the authors have used maximum loops of 50 and a minimum coefficient of determination, called the R-squared value, of 0.99 for the termination criteria. If any of the criteria are satisfied, the identification processes are terminated and the optimized parameters have peaked, according to the minimum of standard error between the data and the model prediction. It should be noted that the proposed criterion of termination here cannot guarantee that a global optimum will be found, but the process is expected to greatly increase the probability of realizing a global optimum, instead of a local optimum.

Based on the parameter identification approach described in this section, the COVID-19 infection dynamics for several countries from North America, South America, Europe, and Asia is found to be captured well by using the generalized logistic function

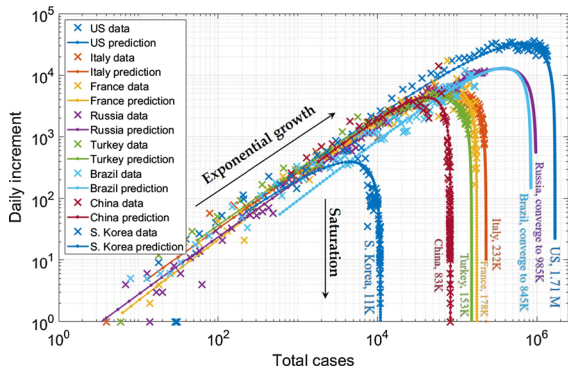


Fig. 4 Phase portraits of COVID-19 dynamics in different countries from North America, South America, Europe, and Asia

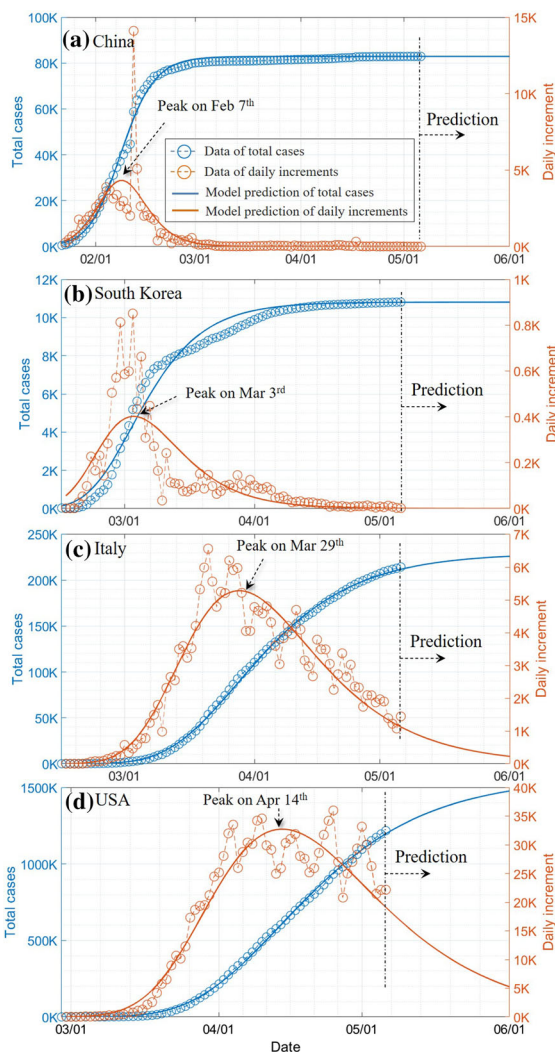


Fig. 5 Generalized logistic function based COVID-19 dynamics: **a** China, **b** South Korea, **c** Italy, and **d** the USA

as shown in Fig. 4. In this figure, the abscissa is the total number of test-positive cases, while the ordinate is the daily increase in the number of positive cases; that is, the derivative of the total cases. Hence, Fig. 4 can be considered as being representative of the phase portraits of the dynamic system. From this plot with logarithmic coordinates, one can discern two periods of the COVID-19 pandemic: a) the exponential growth period and b) the saturation period.

Besides, the time-domain COVID-19 dynamics of China, South Korea, Italy, and the USA are illustrated in Fig. 5a–d, respectively. In these figures, the left axis corresponds to the total number of test-positive cases, while the right axis corresponds to the daily increase in the number of positive cases. The results show the generalized logistic function predictions are consistent with the COVID-19 data of China, Italy, and the USA with a R-squared value greater than 0.995; nevertheless, the R-squared value for South Korea data is 0.981; this is indicative of a rather discernible fitting error for the generalized logistic function in this case, as shown in Fig. 5b. This is mainly because the generalized logistic function cannot be used to capture the plateau from March 10 to April 5 in the daily increments data for South Korea. The inflection points of the total cases, that is, the peaks of the daily increments, are identified with good consistency from the data. The identified peak of COVID-19 in China is on February 7, while the peaks of South Korea, Italy, and the USA are 25 days, 51 days, and 67 days later than China, respectively.

Additional model predicted curves for the countries and regions all over the world are shown in Fig. 6. From this figure, one can choose the countries and regions with total confirmed cases more than 20,000 by May 4. It should be noted that the authors have treated each state of the USA as an individual region due to the large infected population in the USA. Based on the parameter identification approach, through the curves in this figure, the authors show the model identified (predicted) daily increment cases from late January 2020 when the pandemic outbreak started in China, to the middle of August 2020. From these plots, it can be seen clearly that there are remarkable time lags among different regions for the outbreaks of the COVID-19 pandemic. These remarkable time lags indicate that simple models with low degrees of freedom (DOF) are not sufficient to capture the global dynamics of the COVID-19 infections from all over the world. To show the distribution of the outbreak dates of the pandemic, histograms for

Fig. 6 Model prediction of daily case increments for the countries and regions all over the world with confirmed COVID-19 case > 20K by May 4, 2020

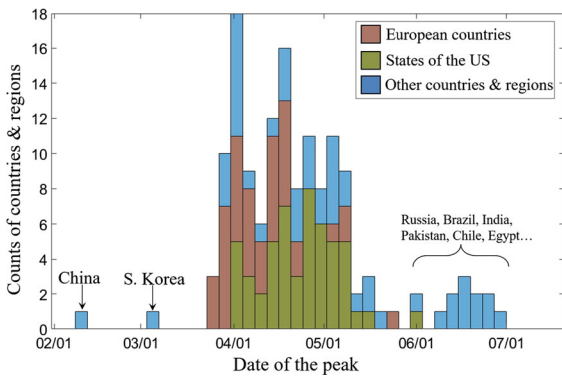
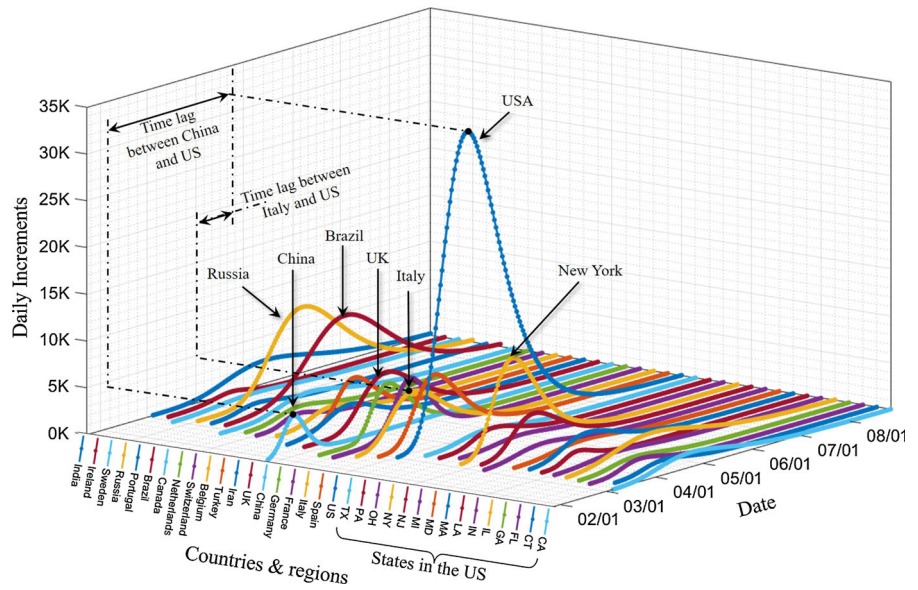


Fig. 7 Histograms for the date distribution of the peaks of COVID-19 for different countries and regions all over the world

the date of peaks identified from the current COVID-19 data for 148 countries and regions all over the world are shown in Fig. 7. This histogram plot is illustrative of the outbreak of the pandemic from one region to another in quick succession after China and South Korea. At the time of writing of the paper, the prediction was that all the countries and regions worldwide would have experienced the peak (at least, the first peak) of the COVID-19 pandemic, by July 1, 2020.

2.3 Global model of COVID-19 dynamics

Given the significant time lags among different countries and regions for the outbreak of the COVID-19

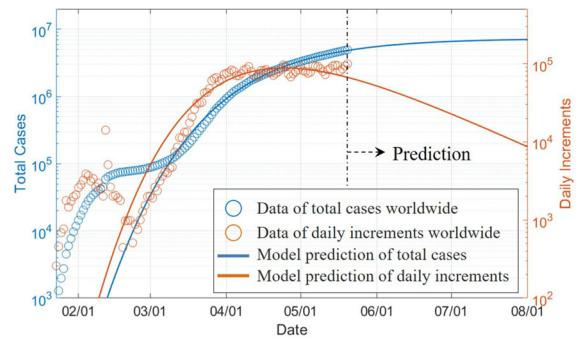


Fig. 8 Dynamics of COVID-19 across the world using single generalized logistic model

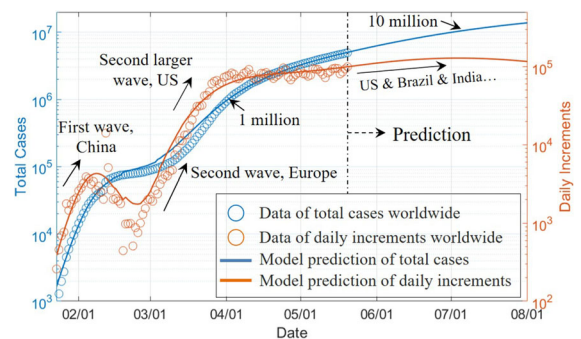


Fig. 9 Dynamics of COVID-19 across the world using the proposed composite global model with 148 identified sub-models

pandemic (Figs. 6, 7), it is natural to expect the global dynamics of the pandemic to be governed by a high-dimensional system; that is, it is not conceivable to capture the global infection dynamics by using a simple model with low degrees of freedom. Hence, inspired by the finite element method (FEM) which is widely used in mechanics, considering each of the regions as an element of the overall global dynamic system, a *composite global model* is established to capture the combined dynamics of countries and regions all over the world. As shown in Fig. 1, the parameter identification approach (Fig. 3) is utilized for each elemental sub-model, following which, the composite global model is constructed by assembling all of the 148 sub-models from different regions all over the world. Thus, the global model for the COVID-19 dynamics at hand is a high-dimensional differential system with 148 sub-models.

The outcome of the COVID-19 dynamics modeling carried out by using a single generalized logistic function is as shown in Fig. 8. The notable deviation between the COVID-19 data and the model prediction with the generalized logistic function confirms the earlier statement that a low-dimensional model cannot be used to capture the global dynamics. By contrast, the outcome of composite global model shown in Fig. 9, which is comprised of 148 identified sub-models, matches the worldwide COVID-19 data with good consistency for both the total number of infection cases and daily increments. The prediction from the model with single generalized logistic function is that the daily increments cases of COVID-19 will drop to 10 thousand by late July; on the other hand, the prediction with the composite global model is that one will have daily increments of more than 100 thousand until August, which is worrisome. With the composite global model, one can see from Fig. 9 that there have been two waves of COVID-19 pandemic spreading across the globe. The first wave started in late January and ended in late February in Wuhan, China. Subsequently, the second wave started mainly in Europe and pushed to a surge into the USA. This surge, which has remained over for a long period in the USA, is being followed by one where the spreading is occurring in Russia, Brazil, and India.

It is important to note that all of the predictions made in this section are based on both the generalized logistic function and the COVID-19 data (positive tested cases). The elemental sub-model of the system is quite simple to avoid any over-fitting of the system dynamics. However, with this data-driven prediction, it is impor-

tant to keep in mind that consideration has not been given to other factors, such as the rising temperature due to the seasonal changes, the differences and coupling between the southern hemisphere and northern hemisphere, measures and policies taken to counter the virus in different regions. To study the effects of these factors on the pandemic, a more sophisticated model needs to be established taking into a range of aspects, including but not limited to different perspectives on infection dynamics, control measures, viral transmission dynamics, stability of the dynamics, long-term predictions, and so forth.

3 Prediction and control of COVID-19 using an improved SEIQR model with distributed time delays

In this section, an improved epidemic model with time-varying parameters and distributed time delays is proposed based on the SEIQR model. With this model, quantitative analysis of control measures and quarantining is conducted. Based on the global COVID-19 data, some key aspects and parameters that are reflective of the effects of the measures and policies taken by different countries are identified and discussed.

3.1 Improved SEIQR model with distributed time delay

The SEIQR model is a compartment model widely used for epidemiological modeling in disease propagation [13,28] as well as computer virus in the internet [29]. Let it be supposed that the total population N is divided into five compartments so that $S(t) + E(t) + I(t) + Q(t) + R(t) \equiv N$, where the state variables $S(t)$, $E(t)$, $I(t)$, $Q(t)$, and $R(t)$ denote the population of ‘Susceptible’ (S), ‘Exposed’ (E), ‘Infectious’ (I), ‘Quarantined’ (Q), ‘Removed’ (R) classes at any time t , respectively, they are defined by follows

- $S(t)$: Susceptible cases. Initially, $S(t)$ equal to the population N , which is assumed to be a constant
- $E(t)$: Exposed cases, infected but not yet infectious, in a latent period
- $I(t)$: Infectious cases, with infectious capacity but neither quarantined nor recovered. *One is considered to be infectious if and only if in this stage*

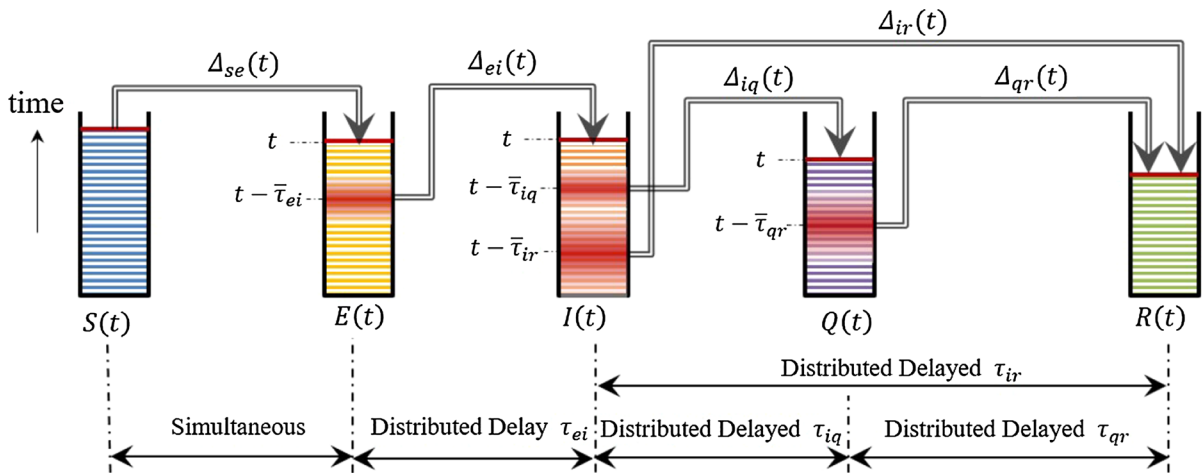


Fig. 10 Construction of improved SEIQR model with delays: flow diagram

- $Q(t)$: Quarantined cases, the infected individuals are confirmed by tests and then isolated from others. Quarantining will lead to loss of infectivity
- $R(t)$: Removed cases (recovered or dead)

As depicted in the flow diagram of Fig. 10, transmissions occur among these four compartments during the COVID-19 pandemic. The dynamics of the transmission can be described as follows. Individuals in the infectious (I) stage can transmit infection to their neighbors through contacts. Individuals in the susceptible (S) stage become infected, and they are assigned to the exposed (E) stage immediately once contacts have occurred. All the exposed (E) individuals are assumed to become infectious (I) after a (random) latency period ranging from 2 to 14 days. The infectious (I) individual might be quarantined (Q) and isolated once they show symptoms and be tested. Finally, quarantined (Q) individuals will be assigned to remove (R) stage after some time. However, according to the recent reports [30, 31], asymptotically infected individuals widely exist, and they can transmit the virus. Consequently, a secondary path of viral transmission needs to be considered; that is, infectious (I) individuals assigned to the removed (R) stage without being quarantined or isolated. It is notable that transmissions among these compartments take some random period for any exposed individual and these time lags during the transmissions result in a multiple distributed delay in the system. Hence, based on the proposed flow diagram of Fig. 10, the dynamical transfer flows among S , E , I , Q , and R

compartments can be written in the form of differential equations as follows:

$$\begin{cases} \dot{S} = -\Delta_{se} \\ \dot{E} = \Delta_{se} - \Delta_{ei} \\ \dot{I} = \Delta_{ei} - \Delta_{iq} - \Delta_{ir} \\ \dot{Q} = \Delta_{iq} - \Delta_{qr} \\ \dot{R} = \Delta_{ir} + \Delta_{qr} \end{cases} \quad (3)$$

here $(\dot{\cdot})$ is the derivative operation with respect to time t . Δ_{jk} is the flow rate from compartment J to compartment K ; that is, the subscripts se , ei , iq , ir , and qr denote the flow rate from S to E , E to I , I to Q , I to R , and Q to R , respectively. Here, the flow rates Δ_{jk} are nonlinear functions, which not only depend on the current state variables but they are also related to the past states of the system. Thus, the time-delay effects are introduced into the system. Intuitively, the time delay here rises from the period of incubation, the period of infection, and the period of quarantine. According to the flow diagram of Fig. 10, the flow rates among the states Δ_{jk} (for $jk = se, ei, iq, ir, qr$) can be written as

$$\begin{cases} \Delta_{se}(t) = \beta \frac{S}{N} I \\ \Delta_{ei}(t) = \int_0^{2\bar{\tau}_{ei}} p_{ei}(\tau) \Delta_{se}(t - \tau) d\tau \\ \Delta_{iq}(t) = \zeta \int_0^{2\bar{\tau}_{iq}} p_{iq}(\tau) \Delta_{ei}(t - \tau) d\tau \\ \Delta_{ir}(t) = (1 - \zeta) \int_0^{2\bar{\tau}_{ir}} p_{ir}(\tau) \Delta_{ei}(t - \tau) d\tau \\ \Delta_{qr}(t) = \int_0^{2\bar{\tau}_{qr}} p_{qr}(\tau) \Delta_{iq}(t - \tau) d\tau \end{cases} \quad (4)$$

Here, β is the infection rate, which is the average number of effective contacts for each infectious individual with others per unit time. β can be a time-dependent variable, that is, $\beta(t)$, and controlled by taking measures against the epidemic. ζ is the quarantine rate, which is assumed to be a constant in this model. Besides these two parameters, there are four distributed delays in this integral equation system Eqs. (4): the delay period from E to I (τ_{ei}), from I to Q (τ_{iq}), I to R (τ_{ir}), and Q to R (τ_{qr}). For different individuals, the period from one state to another can be very different. However, in statistics for a large population, the period of incubation, infection, quarantine, and recovery should follow some form of distribution. This has prompted the authors to consider a novel model with distributed delay in this research. In the literature [32, 33], Weibull distribution, gamma distribution, normal distribution, and other positive skewed distribution are used widely. Without the loss of generality, the authors assume all of the distributed time delays, τ_{jk} (for $jk = ei, iq, ir, qr$), to follow the normal distribution, as shown in Fig. 11. Hence, the probability density function $p_{jk}(\tau_{jk})$ in Eqs. (4) follows the normal distribution with a mean value $\bar{\tau}_{jk}$ and a standard deviation σ_{jk} , and this distribution is written as

$$p_{jk}(\tau_{jk} | \bar{\tau}_{jk}, \sigma_{jk}^2) = \frac{1}{\sigma_{jk}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\tau_{jk}-\bar{\tau}_{jk}}{\sigma_{jk}}\right)^2} \quad (5)$$

for $jk = ei, iq, ir$, and qr

The domain of the probability density function $p_{jk}(\tau_{jk})$ above is all real numbers spanning $-\infty$ to $+\infty$. However, the integral interval of the distributed delay in Eqs. (4) is chosen to be $[0, 2\bar{\tau}_{jk}]$ to not include the non-physical interval less than zero and make the numerical integration feasible. It should be noted that this truncation causes the integration of the density function over the entire integral interval, that is, the area of the shaded region in Fig. 11, to be less than 1. To normalize the probability density function p_{jk} in Eq. (5), a compensation approach can be used by dividing the density function by the shaded area, for instance, dividing p_{jk} by a compensation factor 0.68 for $\sigma_{jk} = \bar{\tau}_{jk}$, by 0.95 for $\sigma_{jk} = \bar{\tau}_{jk}/2$, and by 0.997 for $\sigma_{jk} = \bar{\tau}_{jk}/3$. Here, the authors have chosen $\sigma_{jk} = \bar{\tau}_{jk}/4$ with a compensation factor 0.9999 and the resulting probability density function is shown as Fig. 11.

The proposed, improved SEIQR model at hand, given by Eqs. (3), (4) is a time-varying system with

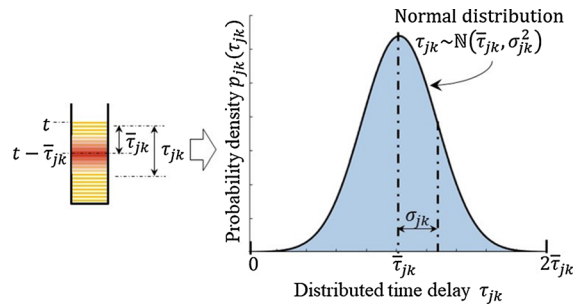


Fig. 11 Construction of improved SEIQR model: the distributed time delay

multiple distributed time delays. Equations (3) form a set of five differential equations, which represents the framework of the SEIQR model. In the first equation in Eqs. (4), Δ_{se} is the transition rate from the compartment of susceptible individuals to the compartment of infectious individuals, and it is also called the force of infection [34]. For a susceptible individual, the transition from S to E is simultaneous once the individual makes an effective contact with an infectious individual. The rest of the four equations in Eqs. (4), which, respectively, represent the flow rate of transition from S to I , I to Q , I to R , and Q to R , are delay integral equations. These four delay integral equations introduce four distributed delays τ_{jk} into the system. The distributed delays τ_{jk} obey normal distribution with mean values $\bar{\tau}_{jk}$ and standard deviations σ_{jk} . Due to the time delay, the system dimension is infinite while the continuity and distribution of the delays introduce the infinite number of time delays into the system. Hence, analytical solutions for the proposed system are impossible and numerical simulation is conducted based on discretization methods [35]. Comparing with the most recent epidemic model with discrete delays given in [36], in the proposed model with distributed delays, one takes the individual differences in symptoms into consideration. This is expected to result in a more realistic dynamic response to the virus and help improve the epidemic model.

3.2 Numerical studies and model comparisons

The proposed dynamical system, which is given by Eqs. (3), (4), is a set of time-varying nonlinear differential equations with multiple distributed delays. Due to the lack of a universal solver for this specific prob-

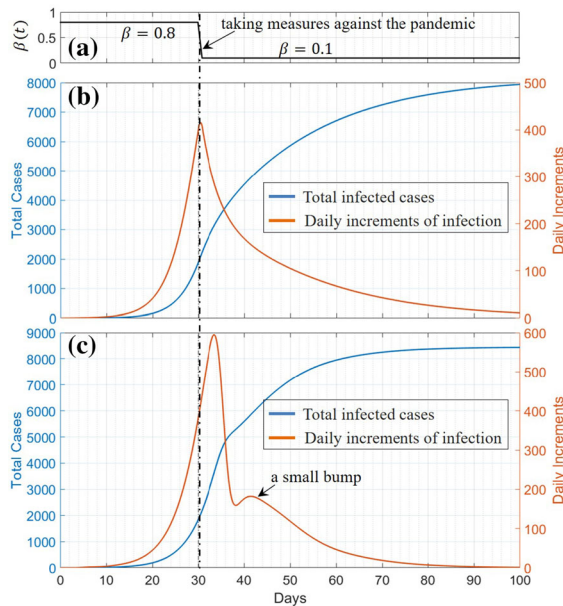


Fig. 12 Comparison of predictions from the conventional SEIQR model and the proposed SEIQR model with distributed delays: **a** time-varying infection rate β , **b** predictions from SEIQR model, and **c** predictions from improved SEIQR model with distributed delays

lem, numerical methods are developed for the simulations and system identification in this research. For the delay integral system described by Eqs. (4), the computations can be done via the convolution algorithm or direct integral. The time step size of 1 day is chosen for the numerical simulations since the historical data of COVID-19 from all the worldwide health organizations are provided in one day intervals. The differential equations given by Eqs. (3) are then integrated via linear four-step Adams–Bashforth methods [37] with the same step size of 1 day. All of the numerical algorithms have been developed based on the MATLAB R2018a platform.

Numerical studies have been conducted with both the conventional SEIQR model and the improved SEIQR model with distributed delays, as shown in Fig. 12. For this comparison, the quarantine rate ζ is set to zero; this degenerates to the SEIR model without quarantine. The population N is set to 10 thousand. The mean values of the distributed delays are set as follows $\bar{\tau}_{ei} = 5$ days, $\bar{\tau}_{iq} = 4$ days, $\bar{\tau}_{qr} = 11$ days, and $\bar{\tau}_{ir} = 15$ days. The standard deviation is set to be one-fourth of the mean values of the distributed delays; that is, $\sigma_{jk} = \bar{\tau}_{jk}/4$. As shown in Fig. 12a, the infection rate

$\beta(t)$ is set to 0.8 initially and drops to 0.1 at the 30th day (the vertical dot-dashed line in the figure), which represents the day on which measures taken to counter the epidemic are put in place. In Figs. 12b, c, the y-axis on the left is the accumulated total infected cases, while the y-axis on the right is the daily increments in infection cases. They can be written as follows:

$$\begin{aligned} \text{Total infected cases} &= \int_0^t \Delta_{ei}(\hat{t})d\hat{t} \\ \text{Daily increments of infection} &= \int_{t-1}^t \Delta_{ei}(\hat{t})d\hat{t} \end{aligned} \tag{6}$$

The dynamic responses obtained with the conventional SEIQR model Fig. 12b and the improved model Fig. 12c are found to be consistent with each other before the measures are taken. However, there are significant differences between the dynamics of these two models after the measures are put in place on the 30th day. In the conventional SEIQR model, an immediate change in the predicted response can be observed with the drop of $\beta(t)$, resulting in a nonsmooth peak. With this model, one notes that the measures taken to counter the virus have an immediate effect on the daily increase in infection cases. Intuitively, this immediate response is not realistic and doesn't agree with the COVID-19 data either. In contrast, with the improved model, the response continues to increase until it reaches a smooth peak for the daily increase in cases 3.8 days later than the date on which measures were initiated. This delayed response is more realistic and conforms with the COVID-19 dynamics worldwide: for instance, in the COVID-19 data in Wuhan, China, one noted a peak of daily increments on February 4, 2020 while the city locked down on January 23, 2020. Besides, with the improved model, one notes a small bump around the fortieth day. This feature matches well the COVID-19 dynamics in many counties and regions, such as South Korea. In summary, when the system is time-invariant, both the conventional SEIQR model and the improved model with distributed delays can capture the dynamics of the epidemic well. When the system is time-varying, which is believed to be the scenario with the COVID-19 pandemic, the improved SEIQR model with distributed delays is more realistic and has significant advantages in capturing the system responses compared to the conventional SEIQR model.

3.3 Quantitative analysis of control measures

Infection control measures to reduce transmission of COVID-19 include universal source control, for example, covering the nose and mouth to contain respiratory secretions, early quarantine, identification, and isolation of patients with suspected disease, use of appropriate personal protective equipment, and environmental disinfection [38, 39]. In the proposed model in Eqs. (3), (4), all the aforementioned control measures are collectively captured through two means: (i) control of infection rate β and (ii) control of quarantine rate ζ .

Quarantine is used to separate someone who has been exposed to COVID-19 and confirmed to be positive by testing, away from others. Quarantine helps prevent the spread of disease by isolating infectious individuals from others. Depending on the policies in different regions, quarantined people may be in isolation or just stay at home. Here, the quarantined individuals are those individuals who have tested positive and are isolated from others. An infectious individual would lose infectivity, once that person is quarantined. The quarantine rate ζ is the ratio of quarantined cases to the infectious cases. Extensive COVID-19 testing and screening can increase the quarantine rate ζ , which requires more coronavirus testing according to the World Health Organization. The quarantine rate ζ and the infection rate β are the only two parameters that the authors can use to control against the spreading of the virus in the improved SEIQR model with distributed time delays, given by Eqs. (3), (4). In this subsection, control of the infection spread is studied by tuning the infection rate β and the quarantine rate ζ .

Two variables, namely “total confirmed cases” and “daily increments of confirmed,” have been introduced into the improved SEIQR model since the states S , E , I , Q , and R in Eq. (3) are almost not realistic to be observed directly in the real world. The total number of confirmed cases and its derivative, that is, the daily increments in confirmed cases, are the most widely used data in epidemiology for COVID-19. According to the proposed model with distributed delays (Fig. 10), they can be written as

$$\begin{aligned} \text{Total confirmed cases} &= \int_0^t \Delta_{iq}(\hat{t})d\hat{t} \\ \text{Daily increments of confirmed} &= \int_{t-1}^t \Delta_{iq}(\hat{t})d\hat{t} \end{aligned} \tag{7}$$

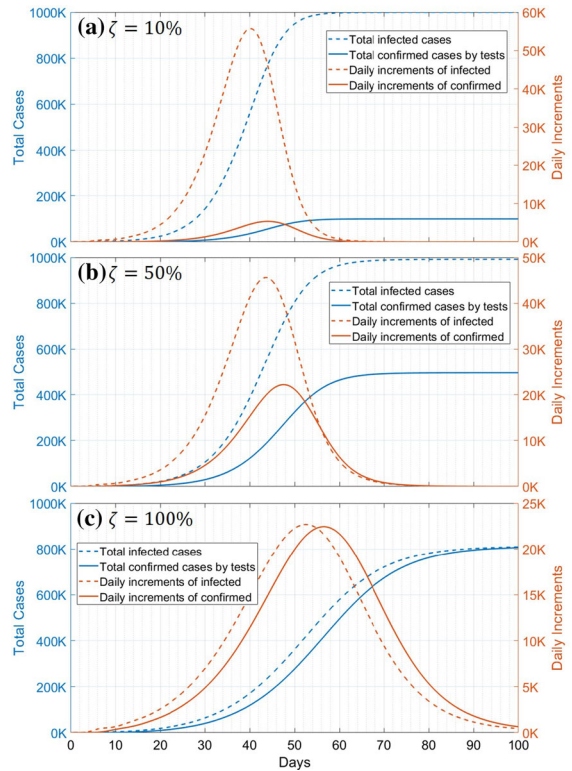


Fig. 13 Control of infection spread through quarantine, for population $N = 1$ million, infection rate $\beta = 0.5$, and different quarantine rates: **a** $\zeta = 0.1$, **b** $\zeta = 0.5$, and **c** $\zeta = 1$

The total cases and daily increments as defined above include only the infectious individuals confirmed by tests and they are observable variables, which distinguishes them from the total infected cases and daily increment of infection, as given in Eq. (6).

To show the effectiveness of quarantining in preventing an epidemic from spreading, quantitative analysis of the control of an epidemic is carried out. As shown in Fig. 13, as the quarantine rate ζ is increased from 10 to 50%, and 100%, the number of total infected cases (dashed blue lines) shows some slight drops from almost 1 million to 0.8 million, while the daily increments of infections (dashed orange lines) show significant drops from 56K per day to 22K per day. However, both the total confirmed cases and daily increments of confirmed are increased due to the increase of COVID-19 testing in quarantine. The results reveal that the total number of infected cases would not have a significant drop by having more quarantining, but the peak of daily infection of an epidemic can be slowed down by increasing the quarantine rate.

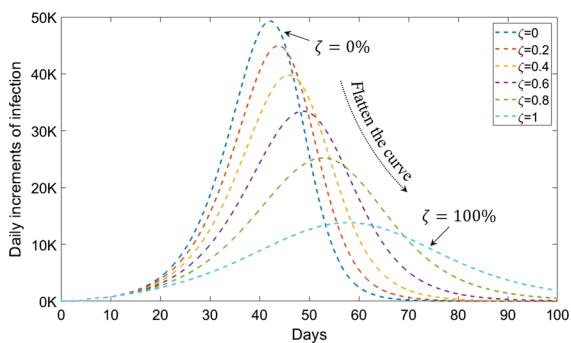


Fig. 14 Flattening of the curve of the COVID-19 epidemic by quarantining, for population $N = 1$ million, infection rate $\beta = 0.4$, and different quarantine rates ζ range from 0 to 1

More comparisons of daily increments of infections for different quarantine rates ζ (0% to 100%) have been carried out, and the corresponding results are shown in Fig. 14. The curves show that an increase in the quarantine rate ζ can “flatten the curve” [40], which can be helpful from the standpoint of a healthcare system. With the effects of “flattening the curve” as illustrated in this figure, the peak value of daily increments of infection reduced from 50K per day to 13K per day, and the peak position is also delayed by 17 days, allowing more time for healthcare capacity to increase and better cope with the patient load. Besides the increase in quarantine rate, decreasing the infection rate β is the other approach to control against the spreading of an epidemic. Effective decrease in the infection rate β is generally due to the control interventions, including lockdown of a city or region, stay-at-home order, social distance measures, wearing masks in public places, and so on. A lot of measures, policies, and orders have been taken by local municipalities and governments and state governments worldwide since the outbreak of the COVID-19. To capture the effects of these measures, the authors have introduced a time-varying infection rate $\beta(t)$ into the COVID-19 model.

As shown in Figs. 15a, c, e, a piecewise linear function is used to describe the time-varying infection rate $\beta(t)$. The drop of $\beta(t)$ on the 26th day is reflective of the measures and orders that are taken to counter the epidemic. Here, it is assumed that it takes seven days for the measures and directives to completely activated; this results in a slope during the linear drop of the infection rate $\beta(t)$ as seen in the figures. In Fig. 15a, b, the infection rate drops from $\beta_0 = 1$ to $\beta_1 = 0.13$ after taking measures against the epidemic. From the pre-

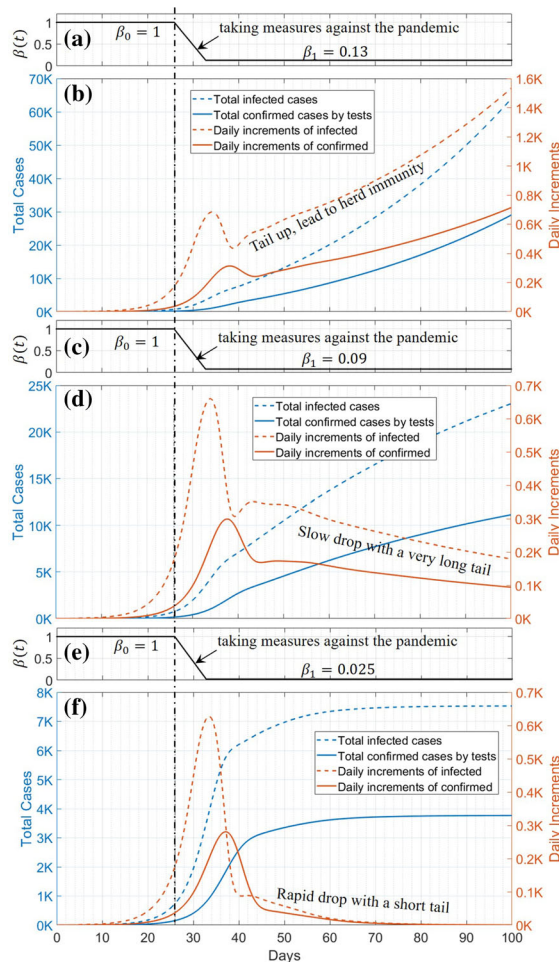


Fig. 15 Control of an epidemic by taking measures, with population $N = 1$ million and quarantine rate $\zeta = 0.5$, **a, c, e** three different time-varying infection rates $\beta(t)$, and **b, d, f** corresponding dynamic responses

dicted system response, it can be observed that the daily increments in infection continuously increases until a local peak is reached 8 days after the measures are initiated. Then, after a short period (five days) of drop, the daily increments of infection start increasing again and the system diverges. A rising tail can overwhelm health systems, leading to high fatalities, and herd immunity eventually. In Fig. 15c, d with the infection rate drop to $\beta_1 = 0.09$, the system response can be said to be convergent and the daily increments of infection drop slowly but with a very long tail. In this scenario, the daily increments of infection are flattened by taking measures, but the total number of infected cases continuously increases almost linearly and this can still

lead to a tremendous number of infected cases after a very long saturation period. In Fig. 15e, f with the infection rate drop to $\beta_1 = 0.025$, the daily increments of infection drop rapidly with a very short tail and there is a quick saturation in the total number of infected cases. For the control of COVID-19, this is the best scenario since the epidemic can be ended within 1 month after taking measures. However, to drop the infection rate drop to 0.025 is a great challenge for both the government and the people.

With the results of Fig. 15b, d, f, the authors have illustrated that slightly different infection rates lead to quite different consequences. An undesirable scenario is one with herd immunity. In this case, one can have break down of healthcare systems and high fatalities. A slightly better scenario is the slow drop case with a very long tail. In this case, the pandemic is under control and one will eventually converge. However, the suffering can be prolonged for a long time. Relatively speaking, the best scenario is the rapid drop situation with a short tail. In this case, the pandemic is ended quickly within about 1 month after taking measures and the suffering is not as prolonged as compared to the previous situation. By examining the current COVID-19 data from all over the world, one can notice that all the COVID-19 curves in different countries and regions can be categorized into these three types mentioned above and the associated dynamics can also be explained by the SEIQR model with distributed delays and time-varying infection rates.

3.4 Data driven dynamics of COVID-19 based on the distributed-delay model

The nonlinear distributed delay systems given by Eqs. (3), (4) provide a feasible mathematical framework to describe the evolution of COVID-19 infection dynamics as well as for modeling the control performances associated with quarantine and other measures. To have a better understanding of the COVID-19 dynamics and to forecast an epidemic’s evolution with high confidence, identification of the nonlinear dynamical system from the time series of COVID-19 data is highly important. Based on the nonlinear regression algorithm and the parameter identification approach used and described in Fig. 3, a similar approach is used to fit the dynamics of the proposed epidemic model Eqs. (3), (4) with the COVID-19 data.

Based on the data from the Web site of The COVID Tracking Project [26] and Worldometers [27], the identified parameters from the data are as follows:

- β_0 : original infection rate before taking any measures
- β_1 : infection rate after taking measures
- t_m : start date for implementation of measures
- ζ : quarantine rate
- R_0 : original reproduction number before taking any measures
- R_1 : reproduction number after taking measures

Here, the reproduction numbers R_0 and R_1 are deduced from the stability of the delay system Eqs. (3), (4), after linearization and discretization of the continuous distributed delay. The reproduction number is a Floquet multiplier for the simplified system of Eqs. (3), (4), and this can be approximated as

$$R_{0,1} = \beta_{0,1}(\zeta \tau_{iq} + (1 - \zeta)\tau_{ir}) \tag{8}$$

where the $R_{0,1}$ is used to denote R_0 or R_1 and $\beta_{0,1}$ is used to denote β_0 or β_1 . The Floquet multiplier $R_{0,1}$ determines the stability of the solution obtained for the epidemic dynamics; that is, if the magnitude of $R_{0,1}$ is greater than 1, the system is unstable and the consequence is herd immunity, which is an undesired scenario as mentioned in the last paragraph of Sect. 3.3. If the magnitude of $R_{0,1}$ is 1, the system is critically stable. If the magnitude of $R_{0,1}$ is less than 1, the system is stable. For the COVID-19 pandemic, the initial reproduction numbers before measures are taken R_0 are always great than 1 (generally between 3 and 4), as identified in Table 1. The reproduction number after taking of measures, R_1 is the key to determine the evolution of epidemic dynamics. The smaller the magnitude of R_1 is, the faster COVID-19 spreading ends.

With the identified parameters from the COVID-19 data as listed in Table 1, the predicted data-driven dynamics of COVID-19 in these countries is shown in Fig. 16. The results illustrate excellent consistency between the COVID-19 data and the predictions from the constructed SEIQR model with distributed delays. The proposed model can capture most of the features of the COVID-19 dynamics observed from data, such as the bump on the daily increments of South Korea from March 9, 2020, to March 25, 2020, and the continuous linear increase in the total infection cases in the USA and UK. The identified date t_m when the measures

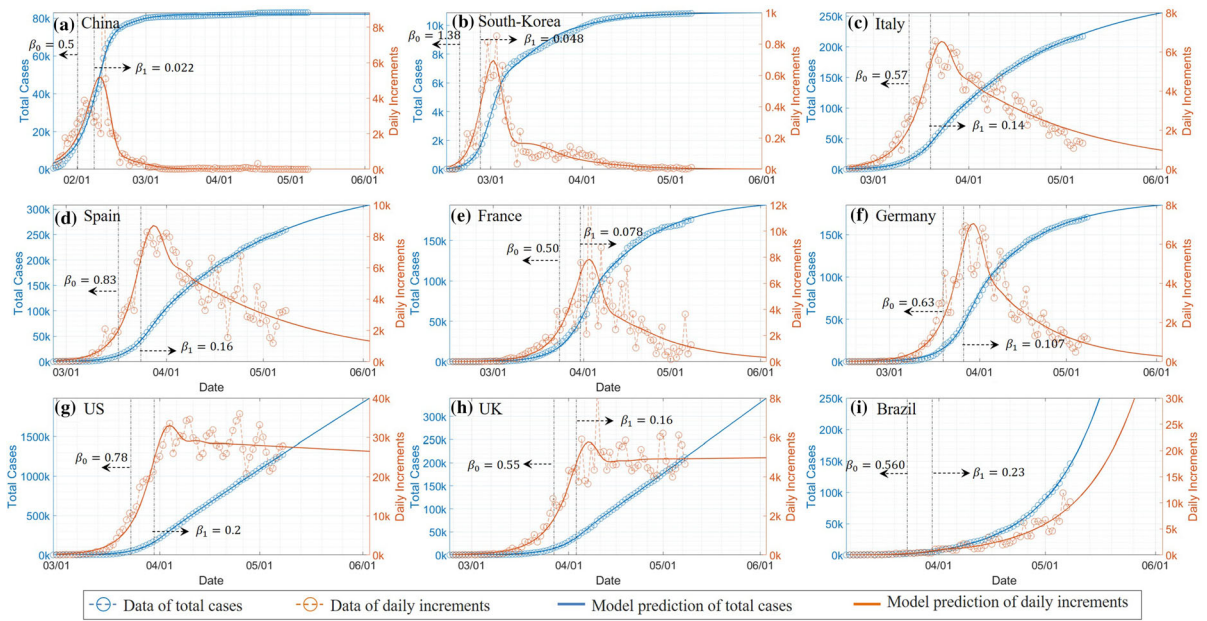


Fig. 16 Identified COVID-19 dynamics from data of different countries (upto May 7 data), based on the model with distributed delays

Table 1 Key parameters identified from the COVID-19 data of different countries worldwide, based on the proposed epidemic model with distributed delays Eqs. (3), (4)

Countries	β_0	β_1	t_m	ζ	R_0	R_1
China	0.50	0.022	Jan. 31	0.65	3.90	0.17
South Korea	1.38	0.048	Feb. 19	0.54	12.5	0.43
Italy	0.57	0.140	Mar. 12	0.85	3.22	0.79
Spain	0.83	0.160	Mar. 17	0.91	4.16	0.80
France	0.49	0.078	Mar. 23	0.73	3.41	0.54
Germany	0.63	0.107	Mar. 19	0.84	3.63	0.62
US	0.78	0.200	Mar. 23	0.92	3.81	0.98
UK	0.55	0.160	Mar. 27	0.78	3.50	0.99
Brazil	0.56	0.230	Mar. 23	0.58	4.82	1.98

were initiated also matches the reality, for example, the identified date t_m of China is January 31, 2020, while the actual period of lock down of cities and stay-at-home order taken by the Chinese government is in the window of January 23, 2020,–January 28, 2020.

From the results shown in Fig. 16 and Table 1, one can also assess the effectiveness of measures taken to counter COVID-19 in different countries. As listed in Table 1, the R_1 numbers of China and South Korea indicate that they these countries undertook measures, which helped best control the COVID-19 spread, with a reproduction number being $R_1 < 0.5$. The effective-

ness of COVID-19 control in France, Germany, Italy, and Spain were effective but with a little higher reproduction number R_1 . Comparatively speaking, the COVID-19 control measures in the USA and UK have not been as effective at the time of writing this paper, as those undertaken in the previously mentioned countries, since the reproduction number R_1 is close to 1; that is, the system has critical stability. In this scenario, the daily infected cases will continue to drop gradually a very long tail and one can reach a tremendously high number of total infected cases, as shown in Fig. 16g, h. These results suggest that USA and UK will experience

COVID-19 spreading for an extended period of time. Switching to the southern hemisphere, the reproduction number R_1 of Brazil is still around 2. This means an exponential outbreak of the COVID-19 pandemic. If this reproduction number is not controlled to come under 1 in the following weeks, the scenario can potentially lead to herd immunity, as shown in Fig. 16i. This is not a welcome situation for healthcare systems.

4 Concluding remarks

In this work, the spreading of COVID-19 among different geographical regions worldwide has been modeled and studied based on the concepts of data-driven dynamical systems. First, the authors have used generalized logistic functions to study the local infection spreading in different regions. Based on a nonlinear regression algorithm, the statistical fit of the generalized logistic model has been optimized in a four-dimensional parameter space and the system dynamics prediction and forecasting is driven by the COVID-19 data. Subsequently, inspired by the notion of the finite element method from mechanics, a composite global model with 148 elements (sub-models for different regions) is established. In this composite global model, each of the regions worldwide is regarded as an element of the overall global system and the global model construction is based on COVID-19 data from all over the world. This construction of a global model based on generalized function based statistical models for local regions, and the use of this model to make predictions and forecasting is one of the original contributions of this work. This methodology may be employed for studies on global spreading of other pandemics as well.

As an extension of the generalized logistic function model, which in essence is a two-compartment model, an extended compartment model is constructed based on the SEIQR model. In this model, both time-varying parameters such as time-varying infection rates and distributed time delays to reflect the differences in individual responses to an infection are introduced. This is another important and original aspect of this work. With this model, one is able to quantitatively assess the effectiveness of different control measures taken such as lock down of regions and quarantining. Again, the methodology employed here may be adopted for studies of other epidemics. Based on the COVID-19 data,

some key parameters, which can reflect the effects of the measures and policies taken in different regions and different countries, have been identified and discussed. Based on the observed data-driven dynamics, the following remarks are made.

- (i) There are significant time lags among different regions, for the outbreak of the COVID-19 pandemic, as shown in Figs. 6 and 7. Based on the current data and current situation, from the generalized logistic model prediction, one can glean that most of the countries and regions worldwide would have passed the infection peak of the COVID-19 by July. For understanding the global dynamics of the COVID-19, the proposed composite global model is attractive compared to a model with low degrees of freedom. The composite global model has helped capture two or more waves of COVID-19 sweeping the globe. The first phase came to notice in late January in Wuhan, China. Subsequently, the second phase started in Europe and moved to the USA. The next spreading phase is expected to be dominated by the dynamics in the USA, Russia, Brazil, and India, as shown in Fig. 9.
- (ii) The improved SEIQR model with time-varying infection rate and distributed delays is found to be better suited for understanding COVID-19 infection dynamics with and without control measures. With this model predictions, one is not only able to capture the effectiveness of the control measures but also anomalies such as the bump seen in the South Korea data. Based on the prediction comparison with available COVID-19 infection data, it is clear that with just a conventional SEIQR model, one is not able to capture the COVID-19 dynamics well. With the new distributed delay model presented here, one can also understand how measures such as quarantining can help observations such as “flattening of the curve” (Fig. 14). It is also shown that slight differences in infection rates can lead to quite different consequences. To control the COVID-19 infection dynamics, the measures taken against the epidemic need to be strict enough to guarantee that the infection rate β_1 is less than a critical value, as shown in Fig. 15.
- (iii) Based on the data-driven COVID-19 dynamics studied with the distributed delay model, it is evident the measures taken in countries such as China

and South Korea were effective in dropping the reproduction number R_1 to be below 0.5. With regard to Europe, in France, Germany, Italy, and Spain, after the measures that were taken, the R_1 value dropped to be in the range from 0.6 to 0.8. In the USA and UK, this reproduction number is close to 1. A reproduction number less than 0.5 is desirable, as it means a swift end to the spreading of the epidemic. In general, a number below 1 indicates system stability. On the other hand, when the reproduction number is close to 1, where one has critical stability, the COVID-19 infection dynamics can persist for an extended period of time and even lead to herd immunity. In the southern hemisphere, for the current data from Brazil, one has a reproduction number around 2, which is highly undesirable.

One needs to temper the observations made in this work, by noting that in the modeling undertaken here, many aspects are not captured such as for example, the seasonal variations in temperature. These additional aspects need to be considered as well for appropriate use of findings from the current work.

Acknowledgements The author from Shanghai Jiao Tong University gratefully acknowledge the support received through the Nature Science Foundation of China, Grant No. 11902194. The senior author from the University of Maryland is thankful for the support provided through the Minta Martin Professorship, and the Tools and Technologies for Pandemics Initiative supported by the Center for Engineering Concepts Development and the Neilom Foundation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Lin, Q., Zhao, S., Gao, D., Lou, Y., Yang, S., Musa, S.S., Wang, M.H., Cai, Y.I., Wang, W., Yang, L.: A conceptual model for the coronavirus disease 2019 (Covid-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* **93**, 211–216 (2020)
- Leung, J.T., Wu, K., Liu, D., Leung, G.M.: First-wave covid-19 transmissibility and severity in china outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet.* **395**(10233), 1382–1393 (2020)
- Shim, E., Tariq, A., Choi, W., Lee, Y., Chowell, G.: Transmission potential and severity of covid-19 in South Korea. *Int. J. Infect. Dis.* **93**, 339–344 (2020)
- Remuzzi, A., Remuzzi, G.: Covid-19 and Italy: What next? *The Lancet.* **395**(10231), 1225–1228 (2020)
- Spina, S., Marrazzo, F., Migliari, M., Stucchi, R., Sforza, A., Fumagalli, R.: The response of milan's emergency medical system to the covid-19 outbreak in italy. *Lancet* **395**(10227), e49–e50 (2020)
- Kucharski, A.J., Russell, T.W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R.M., Sun, F., Jit, M., Munday, J.D., et al.: Early dynamics of transmission and control of covid-19: a mathematical modelling study. *The lancet infectious diseases.* **20**(5), 553–558 (2020)
- Wanto, A., Windarto, A.P., Hartama, D., Parlina, I.: Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density. *Int. J. Inf. Syst. Technol.* **1**(1), 43–54 (2017)
- Zheng, X., Agarwal, V., Liu, X., Balachandran, B.: Nonlinear instabilities and control of drill-string stick-slip vibrations with consideration of state-dependent delay. *J. Sound Vib.* **473**, 115235 (2020)
- LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
- Prentice, R.L., Pyke, R.: Logistic disease incidence models and case-control studies. *Biometrika* **66**(3), 403–411 (1979)
- Richards, F.J.: A flexible growth function for empirical use. *J. Exp. Bot.* **10**(2), 290–300 (1959)
- Wu, K., Darcet, D., Wang, Q., Sornette, D.: Generalized logistic growth modeling of the covid-19 outbreak in 29 provinces in china and in the rest of the world. *arXiv preprint arXiv:2003.05681* (2020)
- Blackwood, J.C., Childs, L.M.: An introduction to compartmental modeling for the budding infection disease modeler. *Lett. Biomath.* **5**, 195–221 (2018)
- Wu, J.T., Leung, K., Leung, G.M.: Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in wuhan, china: a modelling study. *Lancet* **395**(10225), 689–697 (2020)
- Balachandran, B., Kalmar-Nagy, T., Gilsinn, D.E.: *Delay differential equations: recent advances and new directions.* Springer, New York (2009)
- Liu, X., Vljajic, N., Long, X., Meng, G., Balachandran, B.: Nonlinear motions of a flexible rotor with a drill bit: stick-slip and delay effects. *Nonlinear Dyn.* **72**(1), 61–77 (2013)
- Liu, X., Long, X., Zheng, X., Meng, G., Balachandran, B.: Spatial-temporal dynamics of a drill string with complex time-delay effects: bit bounce and stick-slip oscillations. *Int. J. Mech. Sci.* **170**, 105338 (2020)
- Michiels, W., Niculescu, S.-I.: Stability and stabilization of time-delay systems: an eigenvalue-based approach. SIAM, Philadelphia (2007)
- Cooke, K.L., Van Den Driessche, P.: Analysis of an seirs epidemic model with two delays. *J. Math. Biol.* **35**(2), 240–260 (1996)
- Wang, W.: Global behavior of an seirs epidemic model with time delays. *Appl. Math. Lett.* **15**(4), 423–428 (2002)
- Young, L.S., Ruschel, S., Yanchuk, S., Pereira, T.: Consequences of delays and imperfect implementation of isolation in epidemic control. *Sci. Rep.* **9**(1), 1–9 (2019)
- Ma, W., Takeuchi, Y., Hara, T., Beretta, E.: Permanence of an sir epidemic model with distributed time delays. *Tohoku Math. J. Second Ser.* **54**(4), 581–591 (2002)

23. Beretta, E., Takeuchi, Y.: Global stability of an sir epidemic model with time delays. *J. Math. Biol.* **33**(3), 250–260 (1995)
24. McCluskey, C.C.: Complete global stability for an sir epidemic model with delay-distributed or discrete. *Nonlinear Anal. Real World Appl.* **11**(1), 55–59 (2010)
25. Huang, G., Takeuchi, Y., Ma, W., Wei, D.: Global stability for delay sir and seir epidemic models with nonlinear incidence rate. *Bull. Math. Biol.* **72**(5), 1192–1207 (2010)
26. The covid tracking project. <http://covidtracking.com> (2020). Accessed 20 May 2020
27. Covid-19 coronavirus pandemic. www.worldometers.info/coronavirus/ (2020). Accessed 21 May 2020
28. Yan, X., Zou, Y., Li, J.: Optimal quarantine and isolation strategies in epidemics control. *World J. Model. Simul.* **3**(3), 202–211 (2007)
29. Xiao, X., Fu, P., Dou, C., Li, Q., Hu, G., Xia, S.: Design and analysis of SEIQR worm propagation model in mobile internet. *Commun. Nonlinear Sci. Numer. Simul.* **43**, 341–350 (2017)
30. Transmission of covid-19 by asymptomatic cases. <http://www.emro.who.int/health-topics/corona-virus/transmission-of-covid-19-by-asymptomatic-cases.html> (2020). Accessed 1 July 2020
31. Kimball, A., Hatfield, K.M., Arons, M., James, A., Taylor, J., Spicer, K., Bardossy, A.C., Oakley, L.P., Tanwar, S., Chisty, Z., et al.: Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility-king county, washington, March 2020. *Morb. Mortal. Wkly. Rep.* **69**(13), 377 (2020)
32. Virlogeux, V., Fang, V.J., Park, M., Wu, J.T., Cowling, B.J.: Comparison of incubation period distribution of human infections with MERS-CoV in South Korea and Saudi Arabia. *Sci. Rep.* **6**, 35839 (2016)
33. Backer, J.A., Klinkenberg, D., Wallinga, J.: Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, china, 20–28 january 2020. *Eurosurveillance* **25**(5), 2000062 (2020)
34. Compartmental models in epidemiology. https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology (2020). Accessed 20 May 2020
35. Insperger, T., Stépán, G.: Updated semi-discretization method for periodic delay-differential equations with discrete delay. *Int. J. Numer. Methods Eng.* **61**(1), 117–141 (2004)
36. Vyasarayani, C.P., Chatterjee, A.: Complete dimensional collapse in the continuum limit of a delayed seiqr network model with separable distributed infectivity. *arXiv preprint arXiv:2004.12405* (2020)
37. Linear multistep method. https://en.wikipedia.org/wiki/Linear_multistep_method (2020). Accessed 21 May 2020
38. Cdc:healthcare infection prevention and control faqs for covid-19. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-faq.html> (2020). Accessed 20 May 2020
39. Infection control in health care and home settings. <https://www.uptodate.com/> (2020). Accessed 20 May 2020
40. Flatten the curve. https://en.wikipedia.org/wiki/Flatten_the_curve (2020). Accessed 21 May 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.